Accredited SINTA 2 Ranking
Decree of the Director General of Higher Education, Research, and Technology, No. 158/E/KPT/2021
Validity period from Volume 5 Number 2 of 2021 to Volume 10 Number 1 of 2026

Published online at: **http://jurnal.iaii.or.id**

JURNAL RESTI
(Rekayasa Sistem dan Teknologi Informasi)
Vol. 8 No. 1 (2024) 100 - 110 e-ISSN: 2580-0760

Predicting Smart Office Electricity Consumption in Response to Weather Conditions Using Deep Learning

Zikri Wahyuzi[1], Ahmad Luthfi[2], Dhomas Hatta Fudholi[3]
[1,2,3]Magister Informatika, Informatika, Universitas Islam Indonesia, Yogyakarta, Indonesia
[1]zikri.wahyuzi@students.uii.ac.id, [2]ahmad.luthfi@uii.ac.id, [3]hatta.fudholi@fti.uii.ac.id

Abstract

This study investigates the intricate relationship between electricity consumption in smart office environments, temporal elements such as time, and external factors such as weather conditions. Using a data set that encompasses electrical consumption statistics, temporal data, and weather conditions, the research employs preprocessing, visualization, and feature engineering techniques. The predictive model for electric energy usage is constructed using deep learning architectures, including Long Short-Term Memory (LSTM), Bidirectional Long Short-Term Memory (Bi-LSTM), Gated Recurrent Unit (GRU), and Bidirectional Gated Recurrent Unit (Bi-GRU). Evaluation metrics reveal that the LSTM model outperforms others, achieving minimal Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE). The study acknowledges the limitations of the data set, particularly when comparing electricity usage during work hours and outside working hours in a residential context. Future research aims to address these limitations, considering detailed meteorological data, missing data imputation, and real-time applications for broader applicability. The ultimate goal is to develop a predictive model that serves as a valuable tool for improving energy management in smart office settings, optimizing electricity usage, and contributing to long-term firm profitability.

Keywords: *smart office; electricity consumption prediction; weather for load forecasting; deep learning; time series.*

How to Cite: Z. Wahyuzi, Ahmad Luthfi, and Dhomas Hatta Fudholi, "Predicting Smart Office Electricity Consumption in Response to Weather Conditions Using Deep Learning", *J. RESTI (Rekayasa Sist. Teknol. Inf.)* , vol. 8, no. 1, pp. 100 - 110, Feb. 2024.
DOI: https://doi.org/10.29207/resti.v8i1.5530

1. Introduction

The corporation has experienced substantial expansion in its business operations, resulting in a notable influence on its work culture [1]. Consequently, the organization must diligently pursue its adaptation methods to achieve consistent profitability [2]. Improving operational cost efficiency will be a major focus to reduce expenses and maximize net profit [3].

The efficiency of business operational costs is achieved through a decrease in electrical energy consumption in office buildings [4]. According to a survey conducted by MEMR-UNDP, offices rank among the highest energy-consuming buildings, with shopping complexes and hotels the only ones surpassing them in energy consumption, as illustrated in Table 1. Office buildings rely primarily on electricity as their main energy source [5]. Essentially, when the cost of using electrical energy increases, the costs of operations also increase, resulting in decreased organization profitability [6].

Therefore, accurate forecasting of energy usage is vital for businesses. Avoiding unnecessary expenses will result in a reduction in operational costs and an optimization of the supply of electrical energy [7]. Research by Bunn and Farmer [8] indicates that a 1% increase in errors when estimating electrical loads can lead to an annual increase in operational expenses of almost $13 million. This suggests that even small inaccuracies in the forecast can significantly affect the operating costs of a company [9].

Table 1. EBTKE & UNDP Survey Results 2019 [5]

Types of Office Buildings	Energy Consumption Intensity
Large Building	160 kWh/m2/year
Medium-sized Building	202 kWh/m2/year

Forecasting electricity use in office buildings presents substantial difficulties due to a multitude of unpredictable variables that have a significant influence [10]. Prediction error is highly influenced by external factors such as weather patterns, work holidays,

Received: 06-11-2023 | Accepted: 13-02-2024 | Published Online: 18-02-2024

employee behavior, and other similar elements [11]. When exploring the external factors of the prediction of electricity use, the research community agrees on the need for a comprehensive understanding of weather-related influences.

Studies by Chenghong Wang et al. emphasize the critical role of weather-dependent factors in shaping energy consumption patterns [12]. These external influences, often deemed unpredictable, require sophisticated modeling approaches that can assimilate the complex interplay of variables. To address this, emerging methodologies, including machine learning algorithms and deep learning models, have been increasingly used to improve the accuracy of electricity usage predictions [13].

The emergence of the smart office concept is in sync with the prediction of electrical consumption influenced by the weather factors mentioned above [14]. This alignment is essential because the concept of a smart office is driven by the imperative need to design workspaces that take advantage of architecture centered on computing and communication technology [15]. The term "high-level architecture smart office" (Figure 1) in recent advances refers to a framework that delineates the structure and incorporation of intelligent technology in the contemporary office setting [16]. By incorporating components of this idea, smart office architecture produces a productive, comfortable, and secure workspace that enhances employee productivity and operational efficiency [17].

In particular, in this context, historical data successfully stored in the smart office system can be utilized to gain in-depth insights, providing valuable information to refine and optimize predictions of electricity consumption [18].

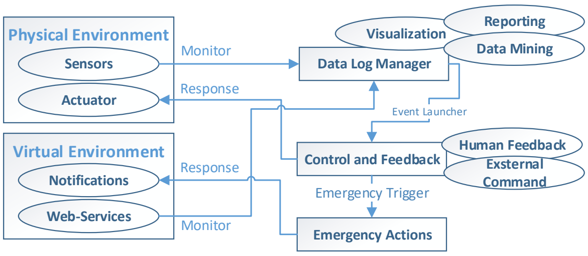

Figure 1. The High-Level Architecture of a Smart Office [16]

Previous research focused primarily on forecasting electricity use by examining intrinsic factors within residential or industrial structures. However, there have only been a limited number of investigations that have examined the use of exterior circumstances in buildings as a factor that can influence power usage [19] - [23].

Consequently, this study aims to gather past data on electricity usage and external weather conditions to examine this particular factor. The previous data will be utilized to forecast office power consumption and offer suggestions for optimizing usage. The novelty of this research lies in its comprehensive exploration of the relationship between electricity consumption, temporal dynamics, and external factors, offering a unique perspective on optimizing energy efficiency in smart office settings.

The study not only identifies existing challenges but also introduces innovative methodologies through the application of advanced deep learning architectures, significantly contributing to the evolving field of smart office energy management. The model is intended to serve as a suggestion system for decision-makers in smart offices, allowing them to supervise, improve, and strategize for more effective energy use. By implementing this approach, organizations can save operational expenses and optimize their earnings.

2. Research Methods

The authors propose a framework to forecast the electrical energy consumption of smart office buildings using four deep learning architectures (LSTM, Bi-LSTM, GRU, and Bi-GRU). Figure 2 shows a flow chart to demonstrate the working approach.

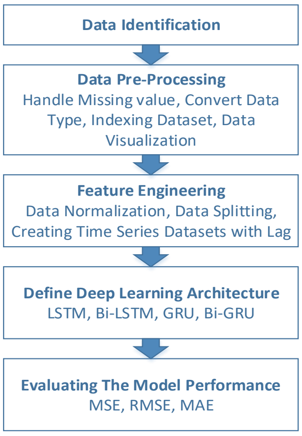

Figure 2. Research Method Flow Chart

The flow chart illustrates the different phases involved in the construction of the suggested prediction model. To achieve greater precision in predicting electricity consumption, especially when taking into account external factors such as weather, it is crucial to have a comprehensive understanding of time series data preprocessing and feature engineering throughout the building of the prediction model [24]. The prediction model construction process can only be launched at that point.

2.1 Data Identification

The author of this study seeks precise data that accurately represent the electricity use of an office building, as well as the corresponding weather conditions in the surrounding area. To achieve this objective, the trials use historical secondary datasets to create predictive models for electricity consumption in smart office systems.

This dataset consists of time series data obtained from an experimental investigation of power usage in a solitary student residence located in the Aarhus Municipality, Denmark [25]. The data source combines historical records of electrical energy usage with weather data acquired from the Danish Meteorological Institute (DMI) [26]. The dataset spans a broad time range, specifically from November 1, 2018, to November 1, 2021, and includes measurements recorded at hourly intervals.

During this period of time, the authors presented meticulous observations of energy consumption, allowing analysis of extended patterns, daily fluctuations, and seasonal variations in electricity consumption over the previous three years. The authors have a valuable dataset that covers a whole year and includes hourly intervals. This information allows them to gather information and create a forecast model for electric energy use. The model is based on external weather factors that affect the office building [27].

The dataset consists of various columns containing different types of information. The "Datetime" column records the date and time, while the "kWh" column records the hourly consumption levels over 3 years. The dataset contains data on the day of the week, month, and whether it is a weekend or not, along with historical weather conditions. Table 2 presents the descriptions, units, and meanings of the variables. The dataset consists of 26,328 rows and 18 columns. To create a model, the date-time variable will be used as an index on the time series dataset. The variable kwh will be the label, while the other columns will serve as features for the prediction model.

Table 2. Dataset Variable Description

Variables	Units	Description
DateTime	-	Information about date and time
kWh	kWh	Amount of energy consumption
hour	hour	Hours of a day
day_of_month	Date	Day of the month
day_of_week	day	Day of the week
month	month	Month
is_weekend	-	Weekend (Saturday & Sunday) or non-weekend
pressure_at_sea	hPa	Atmospheric pressure at sea surface
precip_dur_past10 min	minute	Duration of rainfall in the last 10 minutes
Wind_dir	degree.	Arah angin rata-rata
wind_speed	m/s	Average Wind Speed
temp_dew	°C	Temperatures of dew
pressure.	hPa	Atmosphere Pressure
visib_mean_last10 min	m	Average visibility in 10 minutes
temp_dry	°C	Dry air temperature.
humidity	%	Relative humidity of air
cloud_cover	%	Percentage of sky covered by clouds
visibility	m	Visibility

2.2 Data Pre-processing

Data preprocessing is an essential initial step in data analysis and machine learning. Data preparation is a series of methods to ensure the cleanliness, organization and preparation of data for analysis or modeling [28]. In this regard, the authors describe the data preprocessing techniques used on the electricity consumption dataset to analyze its correlation with meteorological conditions as an external component. During the pre-processing stage before evaluating the developed prediction model, the author used Python 3 programming on the Google Collaboratory platform to simplify the experimentation process.

The initial stage in the data preprocessing phase is to validate the structure of the dataset and acquire descriptive statistics to examine the distribution of values among the variables in the dataset. Once the empty values in the dataset have been identified, they are filled using interpolation techniques [29]. Ensuring that there are no empty values is essential, as they can have a detrimental effect on the integrity of the data and the accuracy of the analysis. To elucidate the linear interpolation technique, consult Formula 1.

$$X_i = X_{i-1} + \frac{(X_{i+1} - (X_{i-1})}{2} \quad (1)$$

X_i denotes the missing valueX_{i-1} represents the value before it, and X_{i+1} indicate the value after it. The application of the linear interpolation technique helps to estimate the missing values based on the pattern of the preceding and subsequent data.

Upon identifying variable values with mismatched data types, a data type amendment will be conducted. The 'DateTime' variable in the dataset is currently of type 'object'. To facilitate data extraction, it will be converted to a date-time data type. The 'DateTime' variable is not employed as a feature or label in the process of constructing the model. Instead, it serves as an indexing dataset to enhance the efficiency of analyzing time-series data processing. It is important to note that time data is still included as a predictive feature in the dataset, including variables for hours, days, and months, as well as weekend and weekday information.

Visualizing the data is the final step in the data pre-processing phase, which is essential for the feature engineering stage. Visualization plays a vital role in examining attributes and structure and uncovering latent data patterns [30].

2.3. Feature Engineering

Feature engineering is the process of developing, altering, or choosing features from a dataset to be used as inputs for a prediction model. The objective is to improve the performance of the model and extract more pertinent information from the available data [31]. Feature engineering encompasses a variety of tasks, including combining features, modifying the volume or shape of the feature distribution, generating new features using additional information, and choosing the most relevant and significant features [32].

The objective of this method is to reduce the number of dimensions in the data, streamline prediction models, and address the problem of unreliable information or excessive noise. Using feature engineering techniques, it is expected that the quality and representation of the data would be improved, leading to more accurate and relevant results from the prediction models [33].

The feature engineering part of this research study involves several activities, such as normalizing the data on the selected features, building a time series dataset by incorporating the lag of those selected features, and ultimately splitting the dataset into testing and training data.

2.3.1 Data Normalization

Normalization is a crucial technique in data processing that aims to standardize the scales of all characteristics, leading to improved model performance [34]. We employed the StandardScaler approach to standardize the data. The StandardScaler method calculates the z score for each feature, which represents the number of standard deviations from the mean. Formula 2 represents the mathematical equation for normalizing data using the Z-score, specifically through the StandardScaler technique.

$$z = \frac{x - \mu}{\sigma} \quad (2)$$

The normalized value (z-score) of the data, denoted as z, is calculated by subtracting the average value μ of the feature from each data point x, and then dividing the result by the standard deviation σ of the feature.

2.3.2 Creating a Time Series Dataset with Lags

The author utilizes the notion of lag generation to create a dataset while creating their time series model. The dataset is created by including previous feature values as inputs, while the target value is projected as future data [35]. Let x_t denote the value of a feature at time t in the time series data. To generate a delay of duration p, the data sequence is arranged according to the structure illustrated in Formula 3.

$$(x_1, x_2, x_3, \dots, x_{t-p}, x_{t-p+1}, \dots, x_{t-2}, x_{t-1}, x_t) \quad (3)$$

The input characteristics utilized for the forecasting of future values consist of $(x_{t-p}, x_{t-p+1}, \dots, x_{t-2}, x_{t-1}, x_t)$, where x_t represents the goal value to be predicted. In this particular case, the value of $p = 24$ is determined using the previous 24 hours of characteristics as input to predict the next kWh value in the data sequence or lag. The goal figure for kWh is derived from the data row that immediately follows the 24-hour sequence. Thoroughly choosing the lag length (p) is of utmost importance, relying on a comprehensive comprehension of the data and prediction goals. Too short of a lag length may not adequately capture significant patterns in the data, whereas an excessively long lag length may mask crucial information regarding trends and variations in the time-series data. Therefore, the determination of the appropriate lag duration is a crucial step in the construction of time series models [36].

2.3.3 Dataset Splitting

Partitioning the dataset into training and testing data is a critical step in the creation of machine learning models. This guarantees the model's capacity to generalize and generate precise predictions on unfamiliar inputs. When there is a dataset with N total samples, separating the datasets involves using a simple mathematical method to distribute a specific proportion [37]. Usually, a training dataset is assigned 70-80% of the total data, while the remaining 20-30% are designated for testing. Formulas 4 and 5 specify the essential mathematical steps to divide the dataset.

$$N_{train} = \ training\ percentage \times N \quad (4)$$

$$N_{test} = \ training\ percentage \times N \quad (5)$$

N_{train} refers to the quantity of training data that has been selected, whereas N_{test} denotes the number of test data that have been chosen. The dataset is divided into two pieces based on a preset ratio using the algorithm given above. This guarantees sufficient data to efficiently train the model and acquire unbiased data to evaluate its performance. Choosing the optimal ratio of training to testing data is essential. Inadequate proportions of training data might result in inadequate learning of the model, resulting in decreased accuracy. On the contrary, if the proportion of data used is too small, the model cannot be adequately tested, which hinders its ability to accurately quantify performance [38].

2.4 Deep Learning Model

The author uses a deep learning approach to examine complex patterns in time series data [39], [40] when forecasting electricity usage in smart office environments. The author employs four types of deep learning architecture in this approach: LSTM (long short-term memory), Bi-LSTM (bidirectional long-short-term memory), GRU (gated recurrent unit), and Bi-GRU (bidirectional gated recurrent unit).

2.4.1 LSTM

LSTM, an abbreviation for long, short-term memory, is a specific kind of Recurrent Neural Network (RNN) that is specifically engineered to preserve and capture long-term dependencies within datasets. LSTMs are highly proficient at capturing intricate and enduring patterns in the prediction of time series [41]. LSTM employs gating mechanisms to retain information for a prolonged duration and discard it when it becomes irrelevant. This effectively solves the problem of the vanishing gradient that is frequently encountered in conventional RNNs. The LSTM model incorporates many types of gates to control the flow of information. The input gate,,, forget gate and output gate are referred to as such [42].

The input gate (i_t) controls the quantity of fresh data stored in the memory cells. The forget gate (f_t) controls the extent to which information from previous memory cells should be retained or discarded. The cell gate (g_t) produces a novel candidate value that is received by the memory cell. The extent to which memory cell values are integrated into the network output is determined by the output gate (o_t).

The updating of memory cells (c_t) is accomplished through the utilization of input, forget, and cell gates. The output (h_t) is obtained by multiplying the value of the memory cell by the gate output. The aforementioned formulas in Figure 3 involve the following variables: x_t represents the input at time t, h_t represents the output at time t, c_t represents the cell value at time t, σ denotes the sigmoid function, tanh denotes the hyperbolic tangent function and represents the element-wise multiplication operation [43].

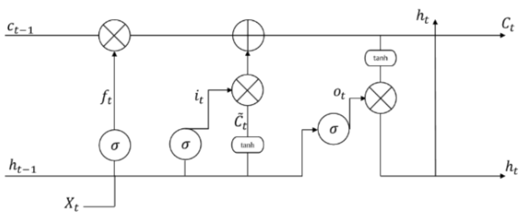

Figure 3. LSTM Architecture

2.4.2 GRU

GRU, or Gated Recurrent Unit, is a form of Recurrent Neural Network (RNN) that shares similarities with LSTM (Long Short-Term Memory) but has a more straightforward architecture. This model is also efficient at addressing time series problems and reduces computing costs compared to LSTM [44]. The GRU provides a harmonious blend of complexity and efficiency. GRU is frequently used when there is a requirement for LSTM performance, but at reduced computational expense. GRU uses gate mechanisms, such as gate reset and gate update, to regulate the flow of information within memory cells [45].

As depicted in Figure 4, The reset gate (r_t) determines the extent to which the information from the previous cell should be deleted. Gate updates (z_t) determine the quantity of fresh information to be stored in memory cells. The candidate value $(\hat{h}_t)$ is computed by utilizing the reset gate (r_t) to merge the information from the input and the previous information $((h_{t-1})$. Next, the value of the memory cell (h_t) is modified by including the gate update (z_t) and the candidate value (ĥ_t) [46].

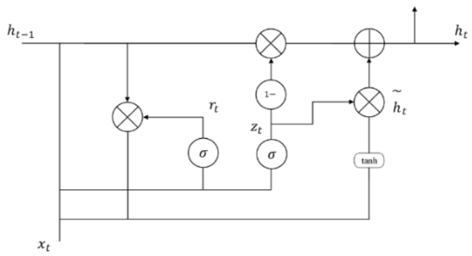

Figure 4. GRU Architecture

2.4.3 Bidirectional RNN (Bi-LSTM & Bi-GRU)

A Bi-RNN, short for Bidirectional Recurrent Neural Network, is a specific kind of neural network design that handles input in two directions: forward (from the beginning to the finish) and backward (from the end to the beginning), as shown in Figure 5. There are two primary categories of Bi-RNNs: Bi-LSTM (Bidirectional Long-Short-Term Memory) and Bi-GRU (Bidirectional Gated Recurrent Unit) [47]. Bi-LSTM is a kind of LSTM that can analyze data from past and future time steps.

The Bi-LSTM model employs two sets of Long Short-Term Memory (LSTM) units: one that processes data in a forward direction and another that processes data in a backward direction. The results of these two Long-Short-Term Memory (LSTM) models are merged to generate the outcome [48]. Bi-GRU is a composite of two GRUs that operate in opposing directions, with one moving forward and the other moving backward. Similarly to Bi-LSTM, the outputs of these two GRUs are merged to obtain the final result [49].

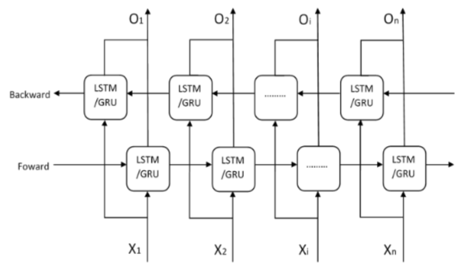

Figure 5. Bi-RNN Architecture

Using Bi-RNN, the model is capable of understanding the intricate connections between these parameters, which allows for more precise and reliable predictions [50]. In the context of this research, Bi-RNN can be utilized to discover energy consumption patterns that are associated with historical weather conditions and future weather forecasts. This can help optimize energy use.

2.5 Model Evaluation

When examining predictive models, the three primary metrics commonly used are the mean squared error (MSE), the root mean squared error (RMSE), and the mean absolute error (MAE). These metrics indicate the proximity between the model's predictions and the actual values in the data. The MSE, RMSE, and MAE metrics offer useful information on model performance. A comprehensive understanding of the merits and limitations of each statistic will facilitate a more precise evaluation of the developed model [51].

The mean squared error is a statistical measure that calculates the average of the squared discrepancies between expected and actual values within a dataset, as shown in Formula 6. A model's predictive accuracy improves as the Mean Squared Error (MSE) decreases. This leads to a more severe penalty for substantial disparities, which guarantees that values that deviate significantly from the true value have a substantial influence on the model evaluation [52].

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (6)$$

The root mean square error (RMSE), as shown in Formula 7, is the square root of the mean squared error (MSE). The term "deviation" refers to the difference between the expected and true values measured in the same units as the variable being measured. RMSE is a more clear and easily accessible computation because it involves taking the square root of MSE [53].

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (7)$$

In contrast, the mean absolute error (MAE) is a statistic that computes the average absolute discrepancy between the projected and actual values. MAE offers insight into the extent of the average discrepancy between the predictions and the actual values. Contrary to the mean squared error (MSE) [54], MAE does not attach greater importance to larger discrepancies between predicted and actual values. Formula 8 represents the mathematical expression for the Mean Squared Error (MSE).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (8)$$

In the three mathematical equations to evaluate the matrix above, n represents the number of samples, y_i represents the true value, and $\hat{y}_i$ represents the predicted value.

3. Results and Discussions

The author uses data analysis and deep learning-based prediction models to forecast the use of smart offices' electrical energy in external weather conditions.

3.1 Smart Office Electricity Consumption Analysis

The energy consumption data set and the external meteorological conditions data will initially be shown to facilitate understanding of the data pattern, thus improving the accuracy of the prediction model. This includes graphical representations of electricity use on an hourly basis (Figure 6), daily basis (Figure 7), weekly basis (Figure 8), and monthly basis (Figure 9).

The graph illustrates that there is a peak in electricity use at the beginning and end of each year, followed by a gradual fall in the middle of the year. Furthermore, when considering the four seasons in the dataset, there is a gradual increase in electricity consumption during the winter months, culminating in its highest point during spring. After spring, the power demand decreases as summer begins and persists until fall.

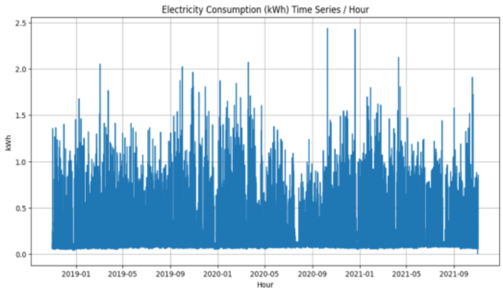

Figure 6. Electricity Consumption / Hour

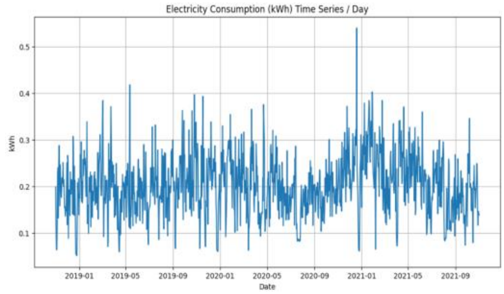

Figure 7. Electricity Consumption / Day

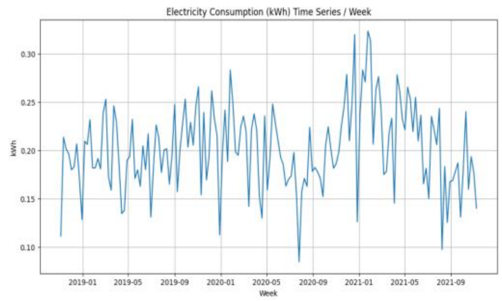

Figure 8. Electricity Consumption/Week

To assess whether electricity use exceeds the monthly average or falls within the usual range, a trend analysis was performed using a moving average (refer to Figure 10).

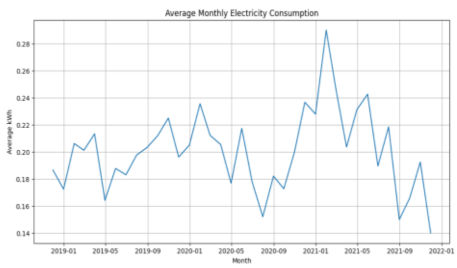

Figure 9. Seasonal Analysis by Plotting Monthly Averages

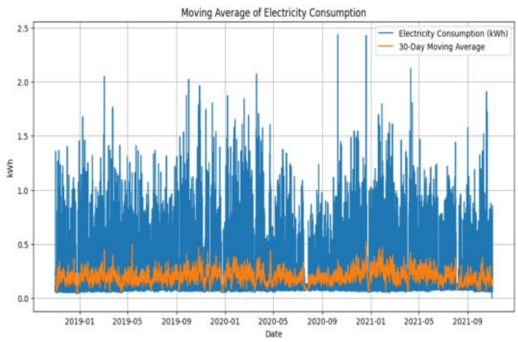

Figure 10. Trend Analysis Using a 30-Day Moving Average

Additionally, a study is conducted to compare electricity consumption patterns on weekdays (Monday–Friday) and weekends (Saturday and Sunday). The findings, depicted in Figures 11 and 12, suggest that electricity use is higher during the week compared to weekends.

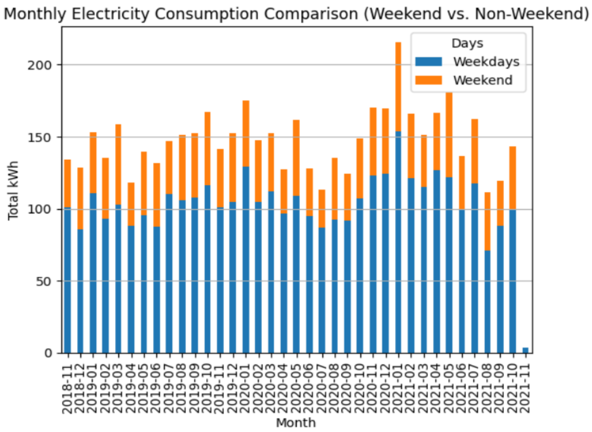

Figure 11. Comparative Analysis of Monthly Average Electricity Consumption on Weekdays & Weekend

Figures 13 and 14 provide a comparative examination of electricity usage during working hours (9:00 am to 5:00 p.m.) and outside of working hours. The statistics indicated that the use of electricity during business hours was lower compared to non-business hours. However, it is crucial to acknowledge the presence of constraints within the data set. It is important to note that the data was collected from studies carried out in a residential setting, rather than an office, and therefore may not provide an accurate representation of energy usage in an office environment.

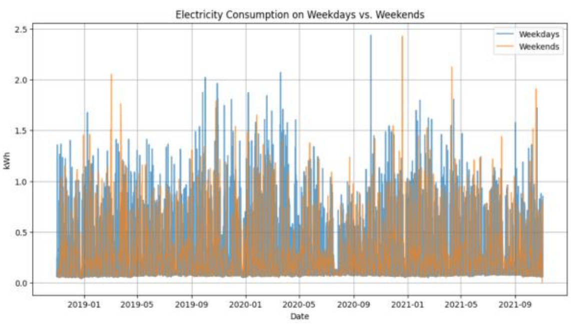

Figure 12. Comparative Analysis of Electricity Consumption on Weekdays & Weekends

Future research should focus on conducting experiments to create an intelligent office system that can accurately measure electricity use during regular working hours (Monday through Friday, 09:00-17:00). The system must have sufficient sophistication to precisely monitor the use of office electricity without the need for human involvement.

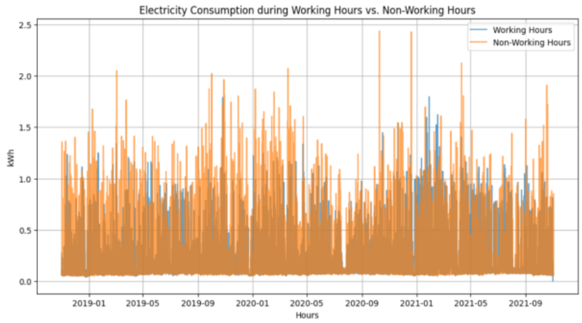

Figure 13. Comparative Analysis of Electricity Consumption During Working Hours 09:00-17:00 & Outside Working Hours

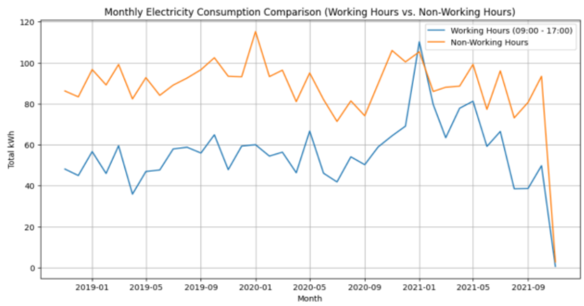

Figure 14. Comparative analysis of monthly average electricity consumption during working hours from 09:00-17:00 & outside working hours

3.2 Smart Office Electric Energy Consumption Prediction Model

When data analysis is performed on the dataset, the subsequent task involves employing a deep learning architecture to construct a system that predicts the usage of electricity. The data analysis findings demonstrate discernible patterns in electrical energy use that exhibit a strong association with external weather conditions and temporal elements that affect office buildings.

The researchers constructed a prediction model using deep learning architectures such as LSTM, Bi-LSTM, GRU, and Bi-GRU, which have a reliable history of properly forecasting time series data. This model effectively forecasts electrical energy usage with a high level of precision. The comparison graph in Figures 15 through 18 illustrates the actual and anticipated electricity usage figures at specific times for the test data sets. Consequently, this graph comparison provides a concise summary of the results obtained from the predictive model for electrical energy consumption in smart office buildings, including external weather and time factors.

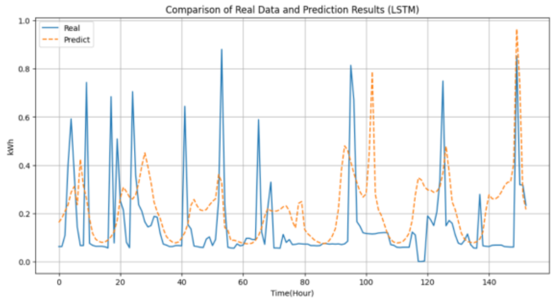

Figure 15. Comparison of Real Data & LSTM Prediction Results

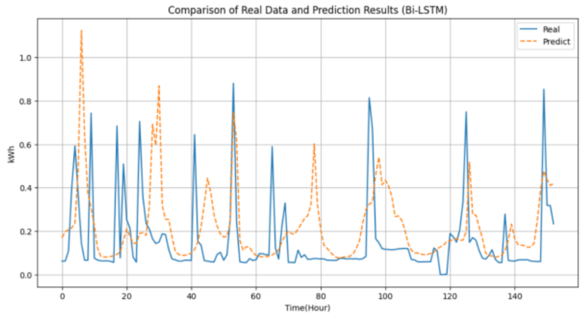

Figure 16. Comparison of Real Data & Bi-LSTM Prediction Results

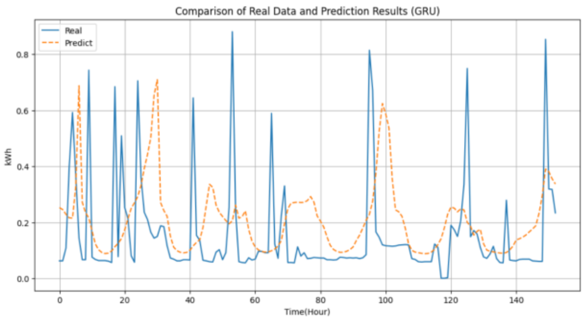

Figure 17. Comparison of Real Data & GRU Prediction Results

When generating an electric energy consumption forecasting model utilizing four distinct deep learning architectures, LSTM, GRU, Bi-LSTM, and Bi-GRU, we also analyze the model performance through analysis of the Mean Squared Error (MSE) graph. The MSE graphs for each of the four architectures provide useful information about the effectiveness of the model in the training data.

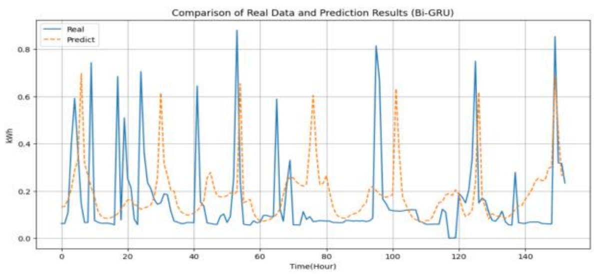

Figure 18. Comparison of Real Data & Bi-GRU Prediction Results

The MSE graph in Figure 19, which shows the LSTM training data, indicates a steady drop in the MSE value from Epoch 1 to Epoch 50, but with a slight increase beyond Epoch 30. However, the graph indicates the reliability of the LSTM model in understanding the patterns of electricity use.

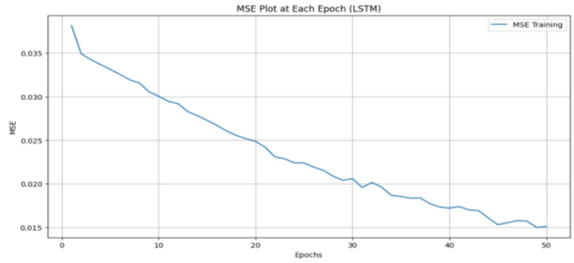

Figure 19. MSE Chart of LSTM Training Data

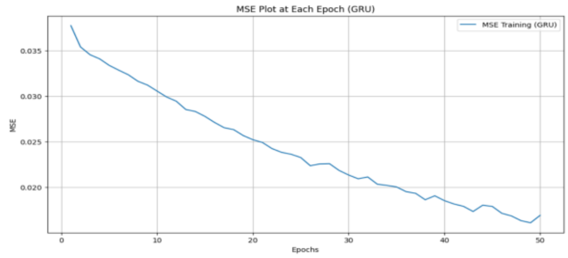

Figure 20. MSE Graph of GRU Training Data

Furthermore, Figure 20's MSE graph exhibiting GRU training data is similarly characterized by a reasonably continuous decline in the number of epochs. However, a slight fluctuation occurred faster than the LSTM at roughly epoch 28, however, the ensuing trend of decline remained obvious. Furthermore, unlike the MSE graph results of the two earlier architectures, the Bi-LSTM exhibited large volatility at the beginning, as shown in Figure 21, especially around epochs 11 and 18, strongly signaling overfitting.

Meanwhile, both the Bi-GRU MSE graph and the LSTM MSE graph exhibit a constant fall, but the Bi-GRU Training Data MSE graph shows a little fluctuation around epoch 28 that is faster than that of LSTM.

However, the pattern of deterioration is continuing, as indicated in Figure 22. Based on the evaluation findings shown in the four MSE graphs, it can be noticed that the LSTM, Bi-GRU and GRU architectures can effectively develop models for predicting electric energy usage. The models show great accuracy and stability in training data, making them a viable alternative to construct an efficient and precise prediction system of electric energy consumption.

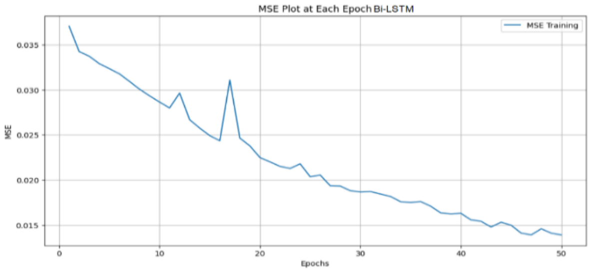

Figure 21. MSE Graph of Bi-LSTM Training Data

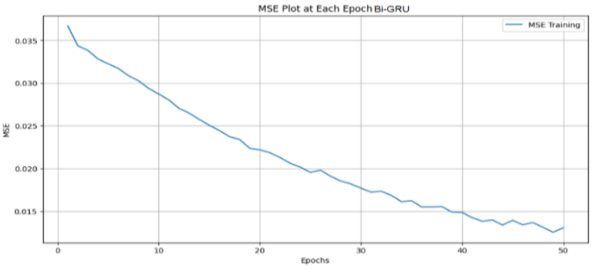

Figure 22. MSE Graph of Bi-GRU Training Data

Based on the findings reported in Table 3, it can be observed that the use of different deep learning architectures provides different outcomes in forecasting the utilization of electrical energy in smart office buildings. However, despite that LSTM presents the lowest MSE value (0.03783) compared to other architectures, it should be noted that this lead is not necessarily proportional to the complexity of the utilized architecture.

With an RMSE score of 0.19450 and an MAE of 0.13871, the LSTM model successfully made extremely accurate predictions. In comparison, the Bi-GRU model produced a remarkable performance with an MSE value of 0.04021, RMSE of 0.20054, and MAE of 0.13184. Despite having a lower complexity than LSTM, Bi-GRU was able to offer accurate forecasts. Similarly, the GRU and Bi-LSTM models also demonstrated good prediction results, although with slightly lower accuracy levels than LSTM and Bi-GRU.

Table 3. Results of the Evaluation Matrix

Evaluation Matrix	LSTM	GRU	Bi-LSTM	Bi-GRU
MSE	0.03783	0.04305	0.04910	0.04021
RMSE	0.19450	0.20749	0.22159	0.20054
MAE	0.13871	0.14242	0.14826	0.13184

4. Conclusions

This study examines the correlation between electricity consumption in smart offices, environmental elements such as time and weather, and energy usage in office buildings. The dataset comprises electrical consumption statistics, temporal data, and weather condition data. The first step consisted of pre-processing and visualizing the data, followed by feature engineering. This involved normalizing the data and generating time series data with lags to enhance accuracy. Construct a predictive model for the use of electrical energy in external factors with the structures of the deep learning model LSTM, Bi-LSTM, GRU, and Bi-GRU.

The development of four prediction models demonstrates the efficient processing of time-series data on office power use using the RNN architecture, taking into account aspects such as time and weather. The matrix assessment results of the four deep learning models demonstrate that the LSTM model outperforms the Bi-LSTM, GRU, and Bi-GRU models. The LSTM prediction model achieves the minimum values for matrix evaluation metrics, namely MSE (0.03783), RMSE (0.19450), and MAE (0.13871). The limitations of the data set hinder the precision of analyzing electricity usage comparisons between working hours and outside of working hours. This is because the dataset is based on a private home experiment rather than actual measurements obtained at an office.

Therefore, the author intends to conduct additional experiments to create an advanced office system that can directly quantify electricity usage during normal office hours (Monday–Friday, 9 a.m.–5 p.m.). The objective of this study is to develop a predictive model for electricity use, which can be used as a valuable suggestion system to improve energy management in smart office settings. Using this methodology, offices can optimize their electricity usage and consequently reduce operational costs, resulting in improved long-term firm profitability.

For improvements in the quality and application of the electric energy consumption prediction model in smart offices, we offer many steps for future research endeavors. First, a more detailed examination of external weather components and time variables can be performed by acquiring comprehensive meteorological data, including metrics such as wind speed, humidity, and precipitation. An in-depth investigation of these aspects will lead to a better understanding of trends in electrical energy consumption. Furthermore, utilizing approaches such as missing data imputation and anomaly detection can strengthen the integrity of the dataset, hence enhancing prediction accuracy. Finally, it is necessary to design real-time applications that take advantage of these predictive models. This program should give users immediate prediction information and enable them to dynamically regulate energy consumption according to the prediction outcomes.

Additionally, it is necessary to test the model in multiple scenarios and locales, taking into account variances in weather conditions and trends in energy consumption. These tests will help identify the benefits and drawbacks of the model in a wider context, ensuring the dependability and relevance of the model in various real-world scenarios. This methodology will enable future research to have a greater impact on achieving energy-efficient and sustainable management in smart office systems.

Acknowledgments

This research was funded by the Ministry of Research and Technology and Higher Education, the Republic of Indonesia, under the Master's Thesis Research Grant scheme, Contract Number: 031/ DirDPPM/ 70 / DPPM/ PPS-PTM-KEMDIKBUDRISTEK/ VI/ 2023.

References

[1] L. D. Parker, "The COVID-19 office in transition : cost , efficiency and the social responsibility business case," vol. 33, no. 8, pp. 1943–1967, 2020, doi: 10.1108/AAAJ-06-2020-4609.

[2] E. Mantesi, K. Chmutina, and C. Goodier, "The office of the future: Operational energy consumption in the post-pandemic era," *Energy Res. Soc. Sci.*, vol. 87, no. June 2021, p. 102472, 2022, doi: 10.1016/j.erss.2021.102472.

[3] S. H. Lin, H. Zhang, J. H. Li, C. Z. Ye, and J. C. Hsieh, "Evaluating smart office buildings from a sustainability perspective: A model of hybrid multi-attribute decision-making," *Technol. Soc.*, vol. 68, no. August 2021, p. 101824, 2022, doi: 10.1016/j.techsoc.2021.101824.

[4] C. E. Danial, A. H. A. Mahmoud, and M. Y. Tawfik, "Methodology for retrofitting energy in existing office buildings using building information modelling programs," *Ain Shams Eng. J.*, vol. 14, no. 6, p. 102175, 2023, doi: 10.1016/j.asej.2023.102175.

[5] Direktorat Jendral Energi Baru Terbarukan Dan Konservasi Energi (EBTKE), "Gedung Perkantoran," Jakarta, 2019. [Online]. Available: https://simebtke.esdm.go.id/sinergi/sektor_pengguna_energi/detail/18/gedung-perkantoran

[6] S. Ali, J. Zhang, A. Azeem, and A. Mahmood, "Impact Of Electricity Consumption On Economic Growth: An Application Of Vector Error Correction Model and Artificial Neural Networks," *J. Dev. Areas*, vol. 54, no. 4, 2020, doi: 10.1353/jda.2020.0039.

[7] O. Jogunola *et al.*, "CBLSTM-AE: A Hybrid Deep Learning Framework for Predicting Energy Consumption," *Energies*, vol. 15, no. 3, pp. 1–16, 2022, doi: 10.3390/en15030810.

[8] D. Bunn and E. D. Farmer, *Comparative models for electrical load forecasting*. 1985. [Online]. Available: https://www.osti.gov/biblio/6256333

[9] G. Chitalia, M. Pipattanasomporn, V. Garg, and S. Rahman, "Robust short-term electrical load forecasting framework for commercial buildings using deep recurrent neural networks," *Appl. Energy*, vol. 278, no. June, p. 115410, 2020, doi: 10.1016/j.apenergy.2020.115410.

[10] S. Chen, Y. Ren, D. Friedrich, Z. Yu, and J. Yu, "Prediction of office building electricity demand using artificial neural network by splitting the time horizon for different occupancy rates," *Energy AI*, vol. 5, no. May, p. 100093, 2021, doi: 10.1016/j.egyai.2021.100093.

[11] J. Hwang, D. Suh, and M. O. Otto, "Forecasting electricity consumption in commercial buildings using a machine learning approach," *Energies*, vol. 13, no. 22, 2020, doi: 10.3390/en13225885.

[12] C. Wang *et al.*, "Impacts of climate change, population growth, and power sector decarbonization on urban building energy use," *Nat. Commun.*, vol. 14, no. 1, pp. 1–16, 2023, doi: 10.1038/s41467-023-41458-5.

[13] G. Vijendar Reddy, L. J. Aitha, C. Poojitha, A. Naga Shreya, D. Krithika Reddy, and G. Sai Meghana, "Electricity Consumption Prediction Using Machine Learning," in *E3S Web of Conferences*, 2023, pp. 1–9. doi: 10.1051/e3sconf/202339101048.

[14] A. Floris, S. Porcu, R. Girau, and L. Atzori, "An iot-based smart building solution for indoor environment management and occupants prediction," *Energies*, vol. 14, no. 10, pp. 1–17, 2021, doi: 10.3390/en14102959.

[15] W. Alomoush, T. A. Khan, M. Nadeem, J. I. Janjua, A. Saeed, and A. Athar, "Residential Power Load Prediction in Smart Cities using Machine Learning Approaches," *2022 Int. Conf. Bus. Anal. Technol. Secur. ICBATS 2022*, no. February, 2022, doi: 10.1109/ICBATS54253.2022.9759024.

[16] A. Barisic, V. Amaral, and M. Challenger, "Enhancing occupants comfort and well-being through a smart office setup," *2020 43rd Int. Conv. Information, Commun. Electron. Technol. MIPRO 2020 - Proc.*, no. September, pp. 1825–1830, 2020, doi: 10.23919/MIPRO48935.2020.9245212.

[17] A. Salosin, O. Gamayunova, and A. Mottaeva, "The effectiveness of the Smart Office system," *J. Phys. Conf. Ser.*, vol. 1614, no. 1, 2020, doi: 10.1088/1742-6596/1614/1/012028.

[18] D. Ramos, P. Faria, A. Morais, and Z. Vale, "ScienceDirect Using decision tree to select forecasting algorithms in distinct electricity consumption context of an office building," *Energy Reports*, vol. 8, pp. 417–422, 2022, doi: 10.1016/j.egyr.2022.01.046.

[19] L. Peng, L. Wang, D. Xia, and Q. Gao, "Effective energy consumption forecasting using empirical wavelet transform and long short-term memory," *Energy*, vol. 238, p. 121756, 2022, doi: 10.1016/j.energy.2021.121756.

[20] M. K. M. Shapi, N. A. Ramli, and L. J. Awalin, "Energy consumption prediction by using machine learning for smart building: Case study in Malaysia," *Dev. Built Environ.*, vol. 5, no. December 2020, p. 100037, 2021, doi: 10.1016/j.dibe.2020.100037.

[21] J. S. Chou, S. C. Hsu, N. T. Ngo, C. W. Lin, and C. C. Tsui, "Hybrid Machine Learning System to Forecast Electricity Consumption of Smart Grid-Based Air Conditioners," *IEEE Syst. J.*, vol. 13, no. 3, pp. 3120–3128, 2019, doi: 10.1109/JSYST.2018.2890524.

[22] B. I. S. Asif Karim, Rootul Patel, "Weather based Electricity Consumption Prediction using Big Data Analytics Model," *Int. J. Eng. Res. Technol.*, vol. 10, no. 4, pp. 158–163, 2021.

[23] Y. Tian, J. Yu, and A. Zhao, "Predictive model of energy consumption for office building by using improved GWO-BP," *Energy Reports*, vol. 6, pp. 620–627, 2020, doi: 10.1016/j.egyr.2020.03.003.

[24] I. H. Sarker, "Machine Learning: Algorithms, Real-World Applications and Research Directions," *SN Comput. Sci.*, vol. 2, no. 3, pp. 1–21, 2021, doi: 10.1007/s42979-021-00592-x.

[25] T. F. GODT, "Personal Electricity Consumption and Weather," *kaggle*, 2021. https://www.kaggle.com/datasets/timongodt/energy-consumption-meteorological-features

[26] DMI Bruger, "Danish Meteorological Institute - Open Data - Meteorological Observation Data - Data Information (metObs) - Parameters (metObs)," *confluence.govcloud.dk*, 2023. https://confluence.govcloud.dk/pages/viewpage.action?pageId=26476616&fbclid=IwAR3IIvZ6_gABMYjlUJ4lwSZX545XY0DEKGIBxVj13utg8Wmuky5w-9DUyS8

[27] D. Mariano-Hernández, L. Hernández-Callejo, F. S. García, O. Duque-Perez, and A. L. Zorita-Lamadrid, "A review of energy consumption forecasting in smart buildings: Methods, input variables, forecasting horizon and metrics," *Appl. Sci.*, vol. 10, no. 23, pp. 1–27, 2020, doi: 10.3390/app10238323.

[28] K. Maharana, S. Mondal, and B. Nemade, "A review: Data pre-processing and data augmentation techniques," *Glob. Transitions Proc.*, vol. 3, no. 1, pp. 91–99, 2022, doi: 10.1016/j.gltp.2022.04.020.

[29] R. Rick and L. Berton, "Energy forecasting model based on CNN-LSTM-AE for many time series with unequal lengths," *Eng. Appl. Artif. Intell.*, vol. 113, no. March, p. 104998, 2022, doi: 10.1016/j.engappai.2022.104998.

[30] A. M. Sharma, "Data Visualization," in *Data Science and Analytics*, S. Kumari, K. K. Tripathy, and V. Kumbhar, Eds., Emerald Publishing Limited, 2020, pp. 1–22. doi: 10.1108/978-1-80043-876-720211002.

[31] D. P. F. Möller, "Machine Learning and Deep Learning," *Adv. Inf. Secur.*, vol. 103, pp. 347–384, 2023, doi: 10.1007/978-3-031-26845-8_8.

[32] P. Duboue, *The Art of Feature Engineering: Essentials for Machine Learning*. 2020. doi: 10.1017/9781108671682.

[33] C. Fan, Y. Sun, Y. Zhao, M. Song, and J. Wang, "Deep learning-based feature engineering methods.pdf," *Appl. Energy*, vol. 240, no. 1, pp. 35–45, 2019, doi: 10.1016/j.apenergy.2019.02.052.

[34] Y. Bengio, Y. Lecun, and G. Hinton, "Deep learning for AI," *Commun. ACM*, vol. 64, no. 7, pp. 58–65, 2021, doi: 10.1145/3448250.

[35] H. Hewamalage, C. Bergmeir, and K. Bandara, "Recurrent Neural Networks for Time Series Forecasting: Current status and future directions," *Int. J. Forecast.*, vol. 37, no. 1, pp. 388–427, 2021, doi: 10.1016/j.ijforecast.2020.06.008.

[36] F. Petropoulos *et al.*, "Forecasting: theory and practice," *Int. J. Forecast.*, vol. 38, no. 3, pp. 705–871, 2022, doi: 10.1016/j.ijforecast.2021.11.001.

[37] S. Raschka, J. Patterson, and C. Nolet, "Machine learning in Python: Main developments and technology trends in data science, machine learning, and artificial intelligence," *Inf.*, vol. 11, no. 4, 2020, doi: 10.3390/info11040193.

[38] L. Alzubaidi *et al.*, *Review of deep learning: concepts, CNN architectures, challenges, applications, future directions*, vol. 8, no. 1. Springer International Publishing, 2021. doi: 10.1186/s40537-021-00444-8.

[39] M. Guven and F. Uysal, "Time Series Forecasting Performance of the Novel Deep Learning Algorithms on Stack Overflow Website Data," *Appl. Sci.*, vol. 13, no. 8, 2023, doi: 10.3390/app13084781.

[40] M. J. Jiménez-Navarro, M. Martínez-Ballesteros, F. Martínez-Álvarez, and G. Asencio-Cortés, "PHILNet: A novel efficient approach for time series forecasting using deep learning," *Inf. Sci. (Ny).*, vol. 632, no. June 2022, pp. 815–832, 2023, doi: 10.1016/j.ins.2023.03.021.

[41] S. Zaheer *et al.*, "A Multi Parameter Forecasting for Stock Time Series Data Using LSTM and Deep Learning Model," *Mathematics*, vol. 11, no. 3, pp. 1–24, 2023, doi: 10.3390/math11030590.

[42] C. Tang, L. Wang, Y. Wei, P. Wu, and H. Wei, "Time-Frequency Domain Variation Analysis and LSTM Forecasting of Regional Visibility in the China Region Based on GSOD Station Data," *Atmosphere (Basel).*, vol. 14, no. 7, 2023, doi: 10.3390/atmos14071072.

[43] B. C. Mateus, M. Mendes, J. T. Farinha, R. Assis, and A. M. Cardoso, "Comparing LSTM and GRU models to predict the condition of a pulp paper press," *Energies*, vol. 14, no. 21, pp. 1–21, 2021, doi: 10.3390/en14216958.

[44] M. Abumohsen, A. Y. Owda, and M. Owda, "Electrical Load Forecasting Using LSTM, GRU, and RNN Algorithms," *Energies*, vol. 16, no. 5, pp. 1–31, 2023, doi: 10.3390/en16052283.

[45] A. Zameer, F. Jaffar, F. Shahid, M. Muneeb, R. Khan, and R. Nasir, "Short-term solar energy forecasting: Integrated computational intelligence of LSTMs and GRU," *PLoS One*, vol. 18, no. 10, p. e0285410, 2023, doi: 10.1371/journal.pone.0285410.

[46] K. Zarzycki and M. Ławryńczuk, "Advanced predictive control for GRU and LSTM networks," *Inf. Sci. (Ny).*, vol. 616, pp. 229–254, 2022, doi: 10.1016/j.ins.2022.10.078.

[47] P. Dey, S. Dev, and B. S. Phelan, "BiLSTM−BiGRU: A Fusion Deep Neural Network For Predicting Air Pollutant Concentration," in *IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium*, 2023, pp. 5166–5169. doi: 10.1109/IGARSS52108.2023.10282742.

[48] P. L. Seabe, C. R. B. Moutsinga, and E. Pindza, "Forecasting Cryptocurrency Prices Using LSTM, GRU, and Bi-Directional LSTM: A Deep Learning Approach," *Fractal Fract.*, vol. 7, no. February 2023, pp. 1–18, 2023, doi: 10.3390/fractalfract7020203.

[49] S. Ahuja, N. A. Shelke, and P. K. Singh, "A deep learning framework using CNN and stacked Bi-GRU for COVID-19 predictions in India," *Signal, Image Video Process.*, vol. 16, pp. 579–586, 2022, doi: 10.1007/s11760-021-01988-1.

[50] N. Putpuek, A. Putpuek, and A. Sungthong, "A Comparative Study of LSTM, GRU, BiLSTM and BiGRU to Predict Dissolved Oxygen," in *2023 20th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, 2023, pp. 67–72. doi: 10.1109/JCSSE58229.2023.10202128.

[51] U. Gupta, V. Bhattacharjee, and P. S. Bishnu, "StockNet—GRU based stock index prediction," *Expert Syst. Appl.*, vol. 207, no. June 2022, p. 117986, 2022, doi: 10.1016/j.eswa.2022.117986.

[52] F. U. M. Ullah, A. Ullah, N. Khan, M. Y. Lee, S. Rho, and S. W. Baik, "Deep Learning-Assisted Short-Term Power Load Forecasting Using Deep Convolutional LSTM and Stacked GRU," *Complexity*, vol. 2022, no. August 2022, pp. 1–15, 2022, doi: 10.1155/2022/2993184.

[53] Y. Feng, Q. Duan, X. Chen, S. S. Yakkali, and J. Wang, "Space cooling energy usage prediction based on utility data for residential buildings using machine learning methods," *Appl. Energy*, vol. 291, no. December 2020, p. 116814, 2021, doi: 10.1016/j.apenergy.2021.116814.

[54] D. Niu, M. Yu, L. Sun, T. Gao, and K. Wang, "Short-term multi-energy load forecasting for integrated energy systems based on CNN-BiGRU optimized by attention mechanism," *Appl. Energy*, vol. 313, no. February, p. 118801, 2022, doi: 10.1016/j.apenergy.2022.118801.

